Structure and nucleotide-induced conformational dynamics of the Chlorobium tepidum Roco protein

Deyaert, Egon; Leemans, Margaux; Singh, Ranjan Kumar; Gallardo, Rodrigo; Steyaert, Jan; Kortholt, Arjan; Lauer, Janelle; Versées, Wim

Published in:
Biochemical Journal

DOI:
10.1042/BCJ20180803

Publication date:
2019

Structure and nucleotide-induced conformational dynamics of the *Chlorobium tepidum* Roco protein

Egon Deyaert[1,2], Margaux Leemans[1,2], Ranjan Kumar Singh[1,2], Rodrigo Gallardo[3,4,#], Jan Steyaert[1,2], Arjan Kortholt[5], Janelle Lauer[6]*, Wim Versées[1,2,]*

[1]VIB-VUB Center for Structural Biology, Pleinlaan 2, 1050 Brussels, Belgium

[2]Structural Biology Brussels, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

[3]Switch Laboratory, Department of Cellular and Molecular Medicine, KU Leuven, Herestraat 49, PB 802, 3000 Leuven, Belgium

[4]VIB Switch Laboratory, VIB, 3000 Leuven, Belgium

[5]Department of Cell Biochemistry, University of Groningen, Groningen 9747 AG, The Netherlands

[6]Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstraße 108, 01307 Dresden, Germany

*Correspondence: Janelle Lauer (jlauer@mpi-cbg.de) or Wim Versées (wim.versees@vub.be).

Current address: Astbury Centre for Structural Molecular Biology, School of Molecular & Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds, LS2 9JT, UK.

Abstract

The LRR-Roc-COR domains are central to the action of nearly all Roco proteins, including the Parkinson's disease-associated protein LRRK2. We previously demonstrated that the Roco protein from *Chlorobium tepidum* (CtRoco) undergoes a dimer-monomer cycle during the GTPase reaction, with the protein being mainly dimeric in the nucleotide-free and GDP-bound states and monomeric in the GTP-bound state. Here, we report a crystal structure of CtRoco in the nucleotide-free state showing for the first time the arrangement of the LRR-Roc-COR. This structure reveals a compact dimeric arrangement and shows an unanticipated intimate interaction between the Roc GTPase domains in the dimer interface, involving residues from the P-loop, the switch II loop, the G4 region and a loop which we named the "Roc dimerization loop". Hydrogen-deuterium exchange coupled to mass spectrometry (HDX-MS) is subsequently used to highlight structural alterations induced by individual steps along the GTPase cycle. The structure and HDX-MS data propose a pathway linking nucleotide binding to monomerization and relaying the conformational changes via the Roc switch II to the LRR and COR domains. Together, this work provides important new insights in the regulation of the Roco proteins.

Introduction

Roco proteins are a class of G proteins that occur in all 3 domains of life, and are characterized by the presence of a Roc (Ras of complex proteins) and a COR (C-terminal of Roc) domain that invariably occur together and are therefore believed to form one functional supra-domain (1–4). While the Roc domain belongs to the superfamily of the Ras-like small GTPases, the function of the COR domain is not entirely known although it has been shown that this domain contributes to the homodimerization of the Roco proteins (5,6). Apart from the Roc and COR domains, Roco proteins can contain several other domains, and based on the domain architecture a classification in three groups has been proposed (1,7). Group 1 Roco proteins display the simplest domain architecture, consisting of the typical Roc-COR domains preceded by a Leucin-Rich Repeat (LRR) domain. Group 2 Roco proteins are characterized by the minimal presence of the LRR, Roc and COR domains, invariably followed by a protein kinase domain plus optional additional domains. Group 3 Roco proteins are characterized by a protein kinase domain that precedes the Roc-COR supradomain and the absence of the LRR domain. In human, 4 Roco protein members are found, distributed over the three groups: MFHAS1 (malignant fibrous histiocytoma amplified sequence 1, previously MASL1) belonging to group 1, LRRK1 and LRRK2 (leucine-rich repeat kinase 1 and 2) belonging to group 2, and DAPK1 (death-associated protein kinase 1) belonging to group 3 (3,8–13).

Research on Roco proteins significantly intensified since 2004, when single point mutations in LRRK2 were linked to autosomal familial Parkinson's disease (PD) (14,15). PD-linked mutations are mainly concentrated in the catalytic Roc-COR and kinase domains of LRRK2. While the clinical PD mutations in the Roc domain decrease the GTPase activity, the most prevalent PD mutation in the kinase domain (G2019S) results in an enhanced kinase activity (16–22).

Despite the huge interest in LRRK2 and related Roco proteins, large gaps are remaining in our understanding of the detailed molecular mechanism of the LRR-Roc-COR domains, common to group 1 and 2 Roco proteins. In 2008 Deng et al. reported a crystal structure of the isolated Roc domain of LRRK2, showing a very unusual domain swapped dimer (23). However, subsequent studies suggest that the domain swapped arrangement likely does not reflect the true conformation of the Roc domain in solution (19). At about the same time

Gotthardt et al. reported the structure of the Roc-COR supradomain of the group 1 Roco protein from the bacterium *Chlorobium tepidum* (5). This structure showed that homodimerization occurs through the COR domains, but resolved only one of the two Roc domains in the dimer. More recently, the crystal structure of the *C. tepidum* Roc-COR supra-domain, together with cross-linking data and low-resolution SAXS and EM data was used to generate a computational model of the 3-dimensional arrangement of a LRRK2 dimer (24).

We discovered that rather than acting as permanent dimers, the Roco proteins undergo a complex dimer-monomer cycle tuned by GTP binding and hydrolysis (25). Concomitant with the GTPase reaction the protein cycles between a compact dimeric state and a seemingly more elongated monomeric state. Moreover, mutations analogous to PD-linked mutations decrease the GTPase activity by interfering with this dimer-monomer cycle. However, the exact nature and mechanism of the nucleotide-induced changes in conformation and oligomerization remained unknown.

Here, we report the first crystal structure showing the arrangement of the LRR-Roc-COR domains of the *C. tepidum* Roco protein in its nucleotide-free form. This structure reveals a compact dimeric arrangement where the LRR domains fold back over the COR domains, and for the first time shows the intimate interactions between the Roc domains in the dimer interface involving residues from the P-loop, the switch II loop, the G4 region and a dedicated loop which we named the “Roc dimerization loop”. The structure moreover proposes a mechanism for linking nucleotide binding to Roco monomerization and intra-subunit domain movements. Hydrogen-deuterium exchange coupled to mass spectrometry (HDX-MS) was then used to map the conformational changes upon GTP binding, GTP hydrolysis and GDP release, and reveal the major events underlying the complex conformational cycle of the Roco proteins. In conclusion, this work provides the first detailed insights in the regulation of the LRR-Roc-COR module underlying the action of Roco proteins.

Experimental Procedures

Protein production, crystallization and data collection

The protocols used for the production of the CtRoc-COR, CtRoco and CtLRR-Roc-COR proteins, and the crystallization and the data collection of the CtLRR-Roc-COR protein have

been described previously (5,25,26). In brief, recombinant protein expression was performed starting from the open reading frames cloned in pProEX plasmids and transformed in *E. coli* BL21 (DE3) cells. All proteins were purified using a three-step protocol consisting of metal chelate affinity chromatography on Ni-NTA, anion exchange chromatography on a source Q30 column and size exclusion chromatography on a Superdex 200 column. Before setting up the crystals the protein was made nucleotide-free by adding an excess of EDTA and loading it on a S200 26/60 gel filtration column equilibrated with 20 mM Hepes pH7.5, 150 mM NaCl, 1 mM DTT and 5% glycerol. Afterwards 5 mM $MgCl_2$ was added to the protein and the nucleotide-load was determined using reversed-phase chromatography coupled to HPLC. Crystals were obtained in 0.8 M Sodium Formate, 0.1 M Sodium Acetate pH 5.5, 10% PEG8000, 10% PEG 1000 and data collection was performed at the PROXIMA2 beamline of the SOLEIL synchrotron (Paris, France) using an EIGER detector (Dectris) (26).

X-ray crystallography data processing and refinement

Data indexing was done using the XDS suite (27) followed by scaling and merging with the AIMLESS program (28). The phase-problem was solved by molecular replacement using the program Phaser and the structures of the individual CtRoc, CtCOR and CtLRR domains as search models (PDB entries 3dpu and 5il7) (5,24,29). The initial model was completed and refined by cycles of manual building in coot and automated refinement using phenix.refine (30).

SEC-SAXS experiment

The protocol used for the SEC-SAXS analysis of the CtLRR-Roc-COR protein has been described previously (24,25). In brief, data was collected on the BM29 beamline at the ESRF in Grenoble (France) using an inline HPLC set-up. 50 µl of an 8 mg/ml sample was injected on a Bio SEC-3 HPLC column (Agilent) equilibrated with 20 mM Hepes pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 5% glycerol and 1mM DTT. Initial processing of the data was done with DATASW (31). The crystallographic model was compared with the experimental SAXS data using CRYSOL (32).

SEC-MALS experiment

SEC-MALS analysis was performed by injecting 10 µl of an 8 mg ml^{-1} protein sample on a Bio SEC-3 HPLC column (Agilent, 3 µm 300 Å) that was equilibrated with 20 mM Hepes pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 5% glycerol and 1 mM DTT. For data collection in the presence of nucleotides, the nucleotide-free proteins were pre-incubated with 1 mM of the nucleotide and 200 µM (= molar excess) of the nucleotide was added to the running buffer. A Dawn Heleos detector (using 9 angles) and Optilab T-rEX detector (Wyatt technology) were attached to the HPLC system (Shimadzu). The molar masses were calculated with the ASTRA 5.3.4.20 software.

Hydrogen Deuterium Exchange Mass Spectrometry

HDX-MS was performed essentially as previously described (33–35). Proteins in 20mM Tris (pH 7.5), 150mM NaCl, 5mM $MgCl_2$, 10% glycerol were diluted 6:4 with 8M urea, 1% trifluoroacetic acid, and passed over an immobilized pepsin column (2.1 mm x 30 mm, ThermoFisher Scientific) in 0.1% trifluoroacetic acid, 5% methanol at 15°C. Peptides were captured on a reversed-phase C8 cartridge, desalted and separated by a Zorbax 300SB-C18 column (Agilent) at 1°C using a 5-40% acetonitrile gradient containing 0.1% formic acid over 12 min and electrosprayed directly into an Orbitrap mass spectrometer (LTQ-Orbitrap XL, ThermoFisher Scientific) with a T-piece split flow setup (1:400). Data were collected in profile mode with source parameters: spray voltage 3.4kV, capillary voltage 40V, tube lens 170V, capillary temperature 170°C. Where applicable, MS/MS CID fragment ions were detected in centroid mode with an AGC target value of 10^4. CID fragmentation was 35% normalized collision energy (NCE) for 30 ms at Q of 0.25. HCD fragmentation NCE was 35eV. Peptides were identified using Mascot (Matrix Science) and manually verified to remove ambiguous peptides. For measurement of deuterium uptake, protein was diluted 1:9 in 20mM Tris (pH 7.5), 150mM NaCl, 5mM $MgCl_2$ prepared in deuterated solvent. Where applicable, 1mM nucleotide was added to the buffer to ensure saturation of the binding site after dilution into deuterated buffer. Samples were incubated for 0, 60 or 900 sec at 22°C followed by the aforementioned digestion, desalting, separation and mass spectrometry steps. The intensity weighted average m/z value of a peptide's isotopic envelope is compared plus and minus deuteration using the HDX workbench software platform (36). Individual peptides are verified by manual inspection. Sequence repetition within the LRR was unexpectedly high. In total, 16 peptides were found in repetition. The HDX workbench software was unable to

distinguish between them and thus assigned each repeated peptide to the first available slot rather than all possible slots and peptide coverage within this region appeared to suffer. However, once taking this into account, the actual peptide coverage was calculated to be 95%. The apparent structure for each isolated peptide, and thus deuterium uptake was similar to the other peptides within each repeated "family". In addition, the deuterium uptake in this region was unaltered by nucleotide binding. Thus, conclusions regarding structural alterations upon nucleotide binding were unaffected by the sequence repetition. Data are visualized using Pymol. Where applicable, deuterium uptake was normalized for back-exchange by comparing deuterium uptake to a sample incubated in 6M urea, 1.2μM pepsin, 10mM NaH_2PO_4 prepared in deuterated buffer for 12-18h at room temperature and processed as indicated above.

Results

The structure of nucleotide-free *C. tepidum* LRR-Roc-COR shows a compact dimeric arrangement

In the nucleotide-free state, the Roco protein from the bacterium *C. tepidum* (CtRoco) occurs as a homodimer in solution (25), where each subunit is 1102 amino acids long and consists of a LRR, Roc and COR domain followed by a C-terminal stretch of 156 amino acids (Supplementary Fig. S1). While the presence of the LRR, Roc and COR domains is generally conserved among group 1 and group 2 Roco proteins, the C-terminal domain is highly variable among different orthologues. To gain insight into the structure and function of the hallmark LRR-Roc-COR domains, we previously reported the purification and crystallization of a domain construct of CtRoco consisting of amino acids 1 – 946 (CtLRR-Roc-COR), and collected X-ray diffraction data to 3.3 Å resolution (26). Here, the structure was solved using molecular replacement and was refined to 3.3 Å and a R_{work} of 0.23 and R_{free} of 0.28 (Table 1).

Two peptide chains are found in the asymmetric unit, revealing in both copies the presence of the LRR domain (residues 2-410), a linker region between the LRR and the Roc domains (residues 411-450), the Roc domain (residues 451-618), a short linker between the Roc and the COR domains (residues 619-629) and the COR domain (residues 630-940). As previously suggested, the COR domain is composed of two subdomains, which we will further refer to

as N-COR (residues 630-775) and C-COR (residues 796-940), linked by a flexible peptide stretch (residues 776-795). Both peptide chains could be nearly completely traced and modelled except for 4 regions in each chain that could not be modelled due to lack of electron density: residues 476-485, 601-605, 733-744 and 786-789 in chain A, and residues 474-486, 601-605, 729-747 and 792-797 in chain B (Supplementary Fig. S2).

The two protomers present in the asymmetric unit correspond to the presumptive biological dimer that is also present in solution in the nucleotide-free state (5,25). Overall, the protein dimer adopts a compact arrangement, where the LRR domains fold back upon the Roc-COR domains within the same protomer (Fig. 1A). Comparison of the theoretical SAXS curve calculated from our structure with the experimental SAXS data of the CtLRR-Roc-COR protein in solution in absence of any nucleotides (24) yields a Chi^2-value of 1.16 (Fig. 1B) (32). This very good match between the experimental and theoretical scattering curves indicates that the compact arrangement shown by our crystal structure is a good representation of the major protein conformation in solution.

Within the homodimer the two protomers are nearly identical with an overall r.m.s.d. of 0.589 Å upon superposition, and with only a very small difference in the relative orientation of the C-COR domains with respect to the other domains. The LRR domains adopt the typical horseshoe shape consisting of 17 successive structural repeat motifs (24). This LRR domain is connected to the Roc GTPase domain via a stretch of 40 amino acids (residues 411-450). The first residues of this connecting region are rich in proline and consist of a ^{414}PLxxPPPE^{421} motif, of which Pro414, Leu415, Pro418 and Glu421 are highly conserved among the prokaryotic Roco proteins. The following residues of the linker region are folded in one short and one longer helix. The latter helix interacts with the first three helices of the N-COR domain and was previously shown to be important for the stable expression of the Roc-COR domain construct of the CtRoco protein (5).

The Roc GTPase domains adopt a very similar conformation as previously observed for the Roc domain in the structure of the CtRoc-COR construct (PDB 3DPU) (5). However, in contrast to the latter structure, where only one of the two Roc domains in the homodimer was resolved, both Roc domains are clearly resolved and could be unambiguously built in the current structure. Similar to other small G proteins, the Roc domain adopts a canonical

arrangement with a central 6-stranded β-sheet flanked on both sides by α-helices, and displaying the highly conserved G1 – G5 motifs (37,38). While the P loop (G1, residues 457-466), the switch II region (G3, residues 513-532) and the loop spanning the G4 motif (residues 569-580) could be entirely traced, electron density is missing due to flexibility for a part of the switch I region (G2, residues 470-489) and the loop spanning the G5 motif (residues 597 – 607) (Fig. 1C). Since the latter two regions typically interact with the γ-phosphate and guanine base of the nucleotide, respectively, it is expected that these regions will become structured upon GTP/GDP binding. One marked difference in the conformation of the Roc domain of the current CtLRR-Roc-COR structure in comparison to the previously determined CtRoc-COR structure (3DPU) is found in the P loop. An additional turn is added at the N-terminus of the first α-helix of the Roc domain, corresponding to residues ^{461}GMAG^{464} of the conserved G1 motif ^{459}GxxxxGKT^{466}. As a consequence of this conformational difference, fitting of a nucleotide in the active site of the CtLRR-Roc-COR structure by superposition with GppNHp-bound Ras (39) would result in severe steric clashes between these residues of the P loop and the phosphates of the nucleotide (Fig. 1C). This means that conformational changes of the P loop residues would be required to allow nucleotide binding. Intriguingly, this same region of the P loop is also implicated in the Roc-Roc dimer interface by forming interactions with the same P-loop residues and with residue Arg543 of the 541-551 loop ("*Roc dimerization loop*") of the adjacent subunit (see further).

The Roc GTPase domain is connected to the COR domain via a short stretch of 10 amino acids, that also contains two Pro residues (Pro620 and Pro627) separated by 1 turn of α-helix. The COR domain adopts the same overall fold as previously observed for the CtRoc-COR structure (5), and consists of an N-terminal N-COR subdomain and a C-terminal C-COR subdomain connected by a 20 amino acid long linker that is partially flexible. N-COR is mainly α-helical and consists of two 3-helix bundles and a 3-stranded β-sheet. One relatively long loop (spanning residues 733-744 and 729-747 in the A chain and B chain, respectively) within the second 3-helix bundle could not be modelled due to flexibility. The C-COR domain consists of a mixed 6-stranded β-sheet flanked by 2 α-helices, followed by a C-terminal part consisting of 2 α-helices and a β-hairpin that is involved in dimer formation (see further).

The Roc domains contribute significantly to the Roco dimer interface

While in the previously available structure of the CtRoc-COR dimer only one of the two Roc domains was visible (5), the current structure of CtLRR-Roc-COR reveals both Roc domains, allowing us to study the complete dimer interface in detail and revealing contacts not previously visible.

Analysis with the PISA server shows that the two protein chains in the asymmetric unit of the crystal structure, forming the presumptive biological dimer, bury a total surface interface of 2085 Å² per subunit (Fig. 2A) (40). While it was previously proposed that the COR-COR interaction would be the main determinant for protein dimerization (5), this interaction only buries an interface area of 630 Å² per subunit, accounting for 30% of the total interface area. Our structure shows more elaborate interactions between the Roc domain of one subunit and the Roc and COR domains of the adjacent subunit. These interactions represent an area of 1455 Å² per subunit, accounting for 70% of the total interface area, and many of these interactions prove critical for inter-subunit and inter-domain communication as highlighted by alterations in deuterium uptake during the GTPase cycle (see further).

A first important interface consists of Roc-Roc interactions (Fig. 2A & 2B). Contacts within this interaction interface involve residues from the P loop, the switch II region and a long loop spanning residues 541 – 551 between the 4th β-strand and 3rd α-helix of the Roc domain. Considering its central localization in the dimer interface, we dubbed this region the "Roc dimerization loop" (Fig. 2B & 2C). A central residue from this dimerization loop is Arg543. Although the relatively weak electron density associated with its side chain suggests some flexibility (Supplementary Fig. S2), Arg543 from subunit A forms a salt bridge with Asp515 and is within van der Waals distance to Gly517 of the switch II loop of the adjacent subunit, and it is located close to the region Gly459-Ala463 of the P loop of subunit B. Other residues of the dimerization loop involved in Roc-Roc dimer interactions are Ser546, Asn547, Tyr550 and Trp551 which interact with the same stretch of residues of the dimerization loop of the adjacent subunit. In accordance with its central role in dimer formation, conservation analysis using the Consurf server shows that many residues of the dimerization loop are highly conserved among the Roco proteins (41). Finally, also reciprocal interactions are made between residues from the adjacent P loops.

A second dimerization interface involves interactions between the Roc and the COR domains of the adjacent subunits, involving the N-COR, C-COR and the linker between the latter two subdomains (Fig. 2A & 2B). On the side of the Roc domain these interactions are mainly formed by residues of the loop spanning the G4 motif (Asn574, Pro575, Ser576, Asn578).

The last dimerization interface is formed by interactions between residues of the C-COR domains of both subunits, as previously described (5). This interface contains several residues from the C-terminal β-hairpin and the preceding α-helix of both C-COR domains.

Thus, the current structure shows that the Roc domain is more important in dimer interactions than previously anticipated. The Roc domain forms extensive interactions with the Roc and COR domains of the adjacent subunit, using residues from the P loop, switch II loop, dimerization loop and the region spanning the G4 motif. The latter regions all surround the nucleotide binding pocket and undergo conformational changes upon nucleotide binding. Moreover, in our structure the conformation of the P loop, which is stabilized by dimer interactions between the P loops of both subunits, would hamper nucleotide binding, hence requiring conformational changes upon binding of the nucleotide that in turn could reduce the stability of the dimer interface. These observations could be the basis for the observed shift toward the monomeric state upon nucleotide binding (25).

Switch II is ideally positioned to couple nucleotide binding to changes in domain organization

Previous SAXS and negative stain EM analyses revealed that the CtRoco protein undergoes large intra-subunit conformational changes coupled to nucleotide-induced monomerization. These analyses suggest that upon GTP (analog) binding the CtRoco protein changes from a compact dimer to a more elongated monomeric species. Although a compact dimeric arrangement in the nucleotide-free state agrees with the current crystal structure, no structural information is available regarding the conformational changes that might take place upon nucleotide binding. Studying the CtLRR-Roc-COR structure and individual steps of the GTPase cycle by HDX-MS (see further) now allows us to look in more detail at the inter-domain interactions within each subunit that are affected upon binding of a nucleotide to the Roc domain.

The LRR domain seems to interact only weakly with the Roc-COR domains, although the middle part of the LRR solenoid (9^{th} and 10^{th} repeat) makes some interactions with the C-terminal part of the C-COR domain (Fig. 1A). The LRR domain is connected to the Roc domain by a region of about 40 residues, containing the aforementioned ^{414}PLxxPPPE^{421} motif followed by 2 helices, that could potentially serve as a hinge for conformational changes of the LRR with respect to the Roc domain. Moreover, the C-terminal part of the LRR and the LRR-Roc hinge region also make direct interaction with the Roc domain, mainly involving residues of the C-terminal end of the switch II region of Roc (Fig. 3A). Gln527 and Phe528 of the switch II region are in close contact to Trp399 of the LRR domain, while Phe528 is also within van der Waals distance to Pro420 of the ^{414}PLxxPPPE^{421} motif. Residues from this latter motif moreover interact with Arg532 of the switch II region, with the side chain of Glu421 forming a salt bridge and the main chain carbonyl of Leu415 forming a H-bond with the side chain of Arg532. As such these interactions between the switch II region of Roc and the hinge region linking the LRR and the Roc domain could be responsible for relaying conformational changes upon nucleotide binding to the LRR domain.

The LRR domain and the hinge region connecting the LRR and Roc domains are also in close proximity to the N-COR domain as well as the linker connecting the Roc and N-COR domains, suggesting that a coordinated conformational change of the LRR and COR domains with respect to the Roc domain could take place. Lys384 of the LRR is involved in a salt bridge with the side chain of Asp621 of the Roc-COR linker region, and the main chain carbonyl of Asp621 forms a H-bond with Arg532 of switch II. This places Arg532 of switch II in the centre of an interaction network with the Roc-COR linker (Asp621), the LRR-Roc hinge region (Leu415 and Glu421) and the LRR (Lys384) (Fig. 3A). Glu416 of the ^{414}PLxxPPPE^{421} motif is also within interaction distance of the main chain carbonyl of Pro620 and of Thr626 within the Roc – N-COR linker. Finally, residues of the ^{414}PLxxPPPE^{421} motif and the following residues also make direct interactions with the N-COR domain, with Pro419, Glu421, Ile422, Gln425, Gln432 of the LRR-Roc hinge, interacting with Asp665, Thr671, Tyr675 and Asn678 of the third α-helix of N-COR.

Apart from the central position of the Gln527-Arg532 region of the switch II loop at the junction of the LRR, Roc and N-COR domains, the switch II loop also makes extensive interactions with residues that are located at the interface of the Roc, N-COR and C-COR

domains, as previously described (Fig. 3B) (5,6). Phe516, Gln519, Ile521 and His526 of switch II point toward the interaction interface with the N-COR domain and interact with Asn677, Tyr706, Ser778 and Leu765 of the N-COR. This centralized position of the switch II loop at the dimer interface (Gly515 and Asp517, see before) and at the interface between the Roc, LRR and N-COR domains, make this region ideally suited to couple nucleotide binding to monomerization and inter-domain conformational changes. Moreover, also switch I interacts with the N-COR domain, with Gly486 forming an H-bond with Asn710, and Leu487 being in van der Waals contact distance to Tyr706. Apart from switch I and switch II, the Roc domain also directly interacts via its $\alpha 3$ helix (residues Tyr550, Arg553, His554, Glu556, Lys557, Tyr558, Gly560) with the C-COR domain (residues Asp803, Tyr804, Phe883, Thr884, Asn885) (Fig 3B).

HDX-MS experiments map the nucleotide-induced conformational changes

We next sought to study the transitions between the nucleotide-bound and nucleotide-free (NF) forms by HDX-MS. Full-length CtRoco (1-1102) was incubated with saturating amounts of either GppNHp (as a proxy for GTP) or GDP, and the deuterium uptake profiles were compared with those of the nucleotide-free (NF) protein. In this way we could monitor the changes occurring in the nucleotide exchange and hydrolysis processes.

First, we will detail the structural changes upon binding of GppNHp by NF CtRoco. Upon binding GppNHp the protein undergoes monomerization and further relaxation into a more extended conformation (25). The associated changes observed with HDX-MS can be classified into 4 categories: 1) stabilization due to direct contacts with the GTP analog, 2) destabilization due to monomerization (loss of contacts at the dimer interface), 3) destabilization caused by relaxing into the extended conformation, and 4) unexpected alterations in deuterium uptake (Supplementary Fig. S3). Starting with category 1: the P loop (G1), switch I (G2), switch II (G3), and G4 region, which encompass most of the nucleotide contact sites were all protected upon binding GppNHp (Fig. 4A & 4B). The G5 region, which also presumably contacts the nucleotide, is not visible in the HDX data, but the peptides surrounding G5 show mild protection upon binding GppNHp. Parts of these regions also make up the dimer interface, but the dramatic protection from deuterium exchange seen by nucleotide binding obscures the more mild enhanced exchange one would expect from

monomerization. However, the remaining two main dimerization interfaces show enhanced exchange as one would expect from the loss of direct dimer contacts. This can be seen in the Roc dimerization loop and the adjacent α3 helix (538-565) and the C-COR interface (901-940) (Fig. 4A & 4B). The more interesting results come from categories 3 and 4. It was seen from SAXS and EM data that activation of CtRoco involves not only monomerization, but also a relaxation into a more extended conformation (25). This can be seen in Fig. 4A & 4C as enhanced deuterium exchange in the LRR (371-402) and the hinge region between the LRR and Roc domains (411-451) encompassing the ^{414}PLxxPPPE^{421} motif. HDX-MS allows us to study the full-length protein and illustrates inter-domain communication not fully appreciated by crystallography alone. One can see from the crystal structure that these regions are interfacing with each other as well as with the switch II loop of the Roc domain (Fig. 3), allowing us to speculate how structural alterations in the Roc domain could impact neighboring domains as a sort of relay switch. Our HDX-MS results show that these regions in the LRR domain and the linker between the LRR and Roc domain with the ^{414}PLxxPPPE^{421} motif change conformation and can act as a hinge, allowing CtRoco to shift into a more extended conformation upon monomerization. There are 3 distinct regions falling into category 4, showing unexpected structural alterations upon nucleotide binding. Two of these appear to play a role in transmitting structural alterations between domains. One encompasses 608-655 and 656-674 within the Roc-COR linker and N-COR regions, respectively (Fig. 4). Their position at a domain interface and interacting with the LRR-Roc hinge region suggests that changes within these regions contribute to the structural communication between LRR, Roc, and N-COR domains. Next is 856-883 forming a long α-helix in C-COR, which shows enhanced deuterium uptake and appears to connect the region around switch II and α3 of Roc to the C-COR – C-COR dimerization surface. The last region in category 4 is 706-730 of the N-COR domain, which shows protection upon binding GppNHp (Fig. 4). This region is adjacent to the G4 loop (566-578) in the opposite subunit and might be expected to undergo enhanced exchange upon monomerization. Instead it shows protection from exchange, suggesting it could be generating new contact sites within the same subunit as a means of stabilizing the Roc and/or COR domains.

Moving on to the next phase of the GTPase cycle, we studied the differences in deuterium uptake in GppNHp-bound and GDP-bound CtRoco reflecting changes occurring upon GTP

hydrolysis. In many ways the structure of GDP-bound CtRoco is an intermediate between that of NF and GppNHp-bound protein as much of the nucleotide binding pocket is still occupied and there is not the full dimerization one sees in the NF protein (25). The HDX-MS data reflects this intermediate position in both nucleotide binding and dimerization (Supplementary Fig. S4). One sees the expected alterations in G1-G4, encompassing the known nucleotide interaction sites. Upon loss of interactions with the γ-phosphate, each of these regions is destabilized and shows enhanced deuterium uptake (Fig. 4). Also the shift toward dimerization is apparent as reduced deuterium uptake, or protection of the C-COR dimerization site 901-940 (Fig. 4). One can also see the accompanying return to the compact structure as stabilization in 371-402 and 411-432, of the LRR domain and hinge between LRR and Roc domains (Fig. 4). Many of the regions involved in relaying conformational change between domains, which were discussed above in the context of GppNHp binding, show alterations in deuterium uptake when comparing GppNHp-bound and GDP-bound protein. One example is 608-632, 642-655, and 656-674 of the Roc – N-COR linker region and the 3rd α-helix of N-COR, which together seem to make an interface connecting the LRR hinge, Roc, and N-COR domains, thereby relaying conformational alterations between these domains. These interfaces, created by long-range contacts in the protein, allow communication between distant regions within the protein and are revealed by the HDX-MS data. Another site altered by GDP binding compared with GppNHp binding is 706-724 and 725-730, of the N-COR domain. This region is destabilized by the replacement of GppNHp by GDP (Fig. 4), and is positioned adjacent to switch II. Since it is adjacent to the Roc domain, it could be stabilizing the Roc domain in the monomer state, but is released in the dimer state. Such a release would be coupled to enhanced deuterium uptake as is seen here (Fig. 4). The last region altered by GDP binding is interestingly not altered by GppNHp binding. It is composed of 798-814 and 886-900 within the C-COR domain. This appears to make up an interface connecting the C-COR dimerization surface with switch II (G3) of the Roc domain, and likely helps convey structural alterations between these regions.

Proceeding to the last stage in the GTPase cycle, release of GDP, we studied the differences in deuterium uptake between GDP-bound and NF CtRoco. All of the changes seen here are the logical extension of the aforementioned HDX-MS results one would expect from GDP release and the ensuing shift toward dimerization (Supplementary Fig. S5). One sees

destabilization of 453-506 and 566-592 of the Roc domain, encompassing P loop (G1), switch I (G2), and G4, as expected from release of GDP (Fig. 4). There is no change in switch II (G3), but none was expected since it largely contacts the γ-phosphate. The dimerization loop of Roc (530-565 peptide) and the C-COR dimer interface (911-934 peptide) are stabilized, consistent with a shift towards complete dimerization. The regions of 396-402 (LRR), 411-432, and 439-452 (both hinge between LRR and Roc), all show stabilization or decreased deuterium uptake, which would be caused by the compaction of the structure correlating with dimerization (Fig. 4). Peptides covering the region of 633-673, of the N-COR domain are stabilized and as mentioned above, appear to create an interface between the LRR domain, its hinge, and the Roc domain illustrating communication between these regions (Fig. 4). Lastly, 876-883 of the C-COR domain is stabilized and is apparently interfacing with 551-564 of the Roc domain further illustrating the numerous "relay switches" between domains, which are critical for the dramatic conformational alterations occurring in this protein when it undergoes the multiple steps needed for the GTPase cycle.

A role of Arg543 in Roco dimerization

The current CtLRR-Roc-COR structure, in combination with the HDX data, revealed an important role of the region spanning residues 541-548 in dimerization. Considering its dedicated role in dimerization, we called this loop the "Roc dimerization loop", although it should be emphasized that it is only one of the many structural regions contributing to the dimer interface. Moreover, the Roc α3-helix (residues 550-560) just C-terminal of the Roc dimerization loop interacts with the C-COR. Deuterium uptake shows that both regions undergo conformational changes upon nucleotide binding. One central residue of the Roc dimerization loop is Arg543, which makes interactions with residues of the P loop and switch II regions of the adjacent subunit (Fig. 3C). Interestingly, Arg543 was previously identified as a residue contributing to GTP hydrolysis, and mutation of this residue in the Roco proteins of *C. tepidum* or *Methanosarceri barkeri* was reported to result in a reduced GTPase activity (5,6). Based on this observation it has been proposed that Arg543 could act as an "arginine finger" to complement the catalytic machinery of the adjacent subunit in the GTPase transition state. Our structure suggests an important role of Arg543 in dimer formation. Thus to evaluate the contribution of the Arg543 residue to dimerization we performed size-

exclusion chromatography multi-angle light scattering experiments (SEC-MALS) with the R543A mutant and the wild type (wt) CtRoc-COR protein (a.a. 412-946) in the presence and absence of GDP or GppNHp (Fig. 5). As we reported previously, the wild type CtRoc-COR protein (theoretical $MM_{monomer}$ = 64.5 kDa) behaves as a stable dimer in the nucleotide-free ($MM_{exp.}$ = 115 kDa) and GDP-bound states ($MM_{exp.}$ = 112 kDa) at the conditions and concentrations used in the experiment. Upon addition of GppNHp, the dimer-monomer equilibrium is shifted toward the monomeric form as observed by a larger elution volume on SEC. The MALS data indicate a gradual decrease of the molecular mass throughout the elution peak indicative of a dimer-monomer equilibrium that is mainly shifted toward the monomeric species, with an observed molecular mass of 72 kDa averaged over the elution peak (Fig. 5A). The R543A mutant still behaves as a stable dimer in the nucleotide-free form (MM_{exp} = 116 kDa), while in the GppNHp-bound state a MM_{exp} averaged over the entire peak of 81 kDa is found, again indicative of mainly monomeric protein. However, in contrast to the wild type GDP-bound protein, the elution profile of the GDP-bound R543A CtRoc-COR protein is broader and shifted toward the monomeric species (MM_{exp} = 90 kDa), showing that the R543A mutation destabilizes the dimer interface and shifts the dimer-monomer equilibrium of the protein in the GDP state toward the monomeric form. This disruption of the dimer-monomer cycle, and the concomitant altered domain interactions, could subsequently contribute to the reported reduction in GTPase activity.

Discussion

In this paper we report the first crystal structure of the LRR-Roc-COR module that is central to the action of all group 1 and group 2 Roco proteins, including the PD-associated LRRK2 (1,7). This structure, in combination with HDX-MS results highlighting structural alterations induced by individual steps of the GTPase cycle, gives us much greater insight into the regulation of this class of proteins.

We recently showed that the CtRoco protein undergoes a dimer-monomer cycle during the GTPase reaction, with the protein being mainly dimeric in the nucleotide-free state and monomeric in the GTP-bound state while an intermediate situation occurs in the GDP-bound state (25). The current structure of CtLRR-Roc-COR in the nucleotide-free state shows a

compact homodimer in agreement with our previous EM and SAXS data (25). In contrast to the previously reported structure of the CtRoc-COR construct (5), our structure reveals both Roc domains within the homodimer, demonstrating unanticipated major contributions of Roc-Roc and Roc-COR interactions to the total dimer interface (Fig. 2). One of the most extensive contributions to the Roc-Roc dimerization interface is provided by a loop (residues 541-551) preceding the $\alpha 3$ helix of Roc, which we therefore named the "Roc dimerization loop". One residue of this loop, Arg543, was initially proposed to be a catalytic arginine finger triggering GTP hydrolysis upon dimerization because an R543A mutant shows decreased GTPase activity (5,6). However, we show here that the R543A mutation shifts the equilibrium of GDP-bound protein to the monomeric form compared to the GDP-bound WT protein. This offers an alternative explanation for the reduced activity of R543A within the framework of our model that Roco proteins alternate between a monomeric and dimeric state during the GTPase cycle.

Intriguingly, the CtLRR-Roc-COR structure shows the P loop in an "occluded" conformation which is not compatible with nucleotide binding (Fig. 1C). A similar conformation of the P loop was not observed in the structure of the monomeric *M. barkeri* Roc-CORΔC (lacking the C-COR domain) protein, nor in the CtRoc-COR structure where only one of the two Roc domains was observed (5,6). This suggests that the occluded P loop conformation is stabilized by the Roc-Roc interface. These observations provide a potential mechanistic trigger for the nucleotide-induced Roco monomerization. Within this scenario binding of GTP in the nucleotide binding site would require a conformational change of the P loop and the adjacent switch II loop and Roc dimerization loop that together form the Roc-Roc dimer interface (Fig 1C & Fig 2C), thus disrupting this interface and promoting monomerization. Moreover, the switch II loop is ideally positioned to relay these conformational changes to other parts of the protein, as testified by our HDX-MS results, ultimately coupling protein monomerization to large scale conformational changes within each subunit. Indeed, the switch II loop (513-532), together with the Roc $\alpha 3$ helix (547-559), interacts with the N-COR and C-COR domains of the same subunit (Fig 3B). Additionally, the C-terminal part of the switch II forms an interaction network with residues of the conserved ^{414}PLxxPPPE^{421} hinge motif that links the LRR domain to the Roc domain, as well as with the peptide segment linking the Roc domain to the N-COR domain (Fig 3A). One can easily envision how GTP

binding to the nucleotide binding site could transmit information via switch II to the LRR, N-COR, and C-COR domains, generating dramatic structural alterations such as monomerization and elongation of the monomeric structure. Our HDX-MS experiments confirm the conformational changes in these regions and thus propose a pathway for relaying conformational changes from the nucleotide binding pocket to the LRR and N-COR and ultimately to the C-COR – C-COR dimer interface (Fig. 4).

Mapping of LRRK2 pathogenic and risk variant PD mutations onto the CtLRR-Roc-COR structure reveals that several of these mutations cluster together in regions that are identified as being important for conformational information transfer and are affected by nucleotide binding as assessed by HDX-MS (Supplementary Fig. S6). His554, Tyr558 and Gly559 of the α3 helix of the Roc domain of CtRoco align with Asn1437, Arg1441 and Ala1442 of LRRK2, respectively, while Ser711, Thr721, Tyr804, and Val888 of the CtRoco COR domains align with Val1613, Arg1628, Tyr1699 and Leu1795 of LRRK2. LRRK2 R1441C/G/H and Y1699C are the most common PD mutations within the Roc-COR domains, while the N1437H, A1442P, V1613A, R1628P and L1795F mutations have been associated with PD (42). In our CtLRR-Roc-COR structure His554, Tyr558, Gly559 and Tyr804 cluster together at the interface of the C-COR and the α3 helix and switch II regions of the Roc domain, and all but Tyr804 display altered deuterium uptake upon binding GppNHp (Fig. 4). The previously reported LRRK2 R1398H protective mutant that shows increased GTPase activity aligns with CtRoco Gln519 (43). Interestingly, Gln519 is located within the Switch II region, relatively close to the Roc α3 helix of the same subunit and to the dimer interface (Fig. 3B). The Gln519 side chain points towards Ser778 of the linker connecting the N-COR to the C-COR domain. Also Glu580 locates close to the CtRoco dimer interface at the C-terminal end of the G4 region, and aligns with Lys1468 in LRRK2, of which the K1468E mutation has been identified as a potentially pathogenic LRRK2 variant. Val489 is the homologue of Ile1371 in LRRK2, and a I1371V mutation has also been associated with PD. Val489 is located within the switch I region at the interface with the N-COR domain and switch II and stacks to His699 of N-COR and Trp514 of switch II. There is also one possibly pathogenic LRRK2 mutation (R1325Q) that locates in the hinge region connecting the LRR and Roc domains, which appears to flex causing the elongation associated with monomerization, where Arg1325 corresponds to Lys443 of CtRoco. Finally, a number of other potential pathogenic mutations

are found in the LRR domain. Two of these residues Arg1067 and Ser1096 seem to align with respectively Gly160 and Thr186 in CtRoco, which are located in a region of the LRR that is close to the C-terminal extreme of the C-COR involved in dimer formation. Together, this indicates that the regions that we find important here are also key to the function of LRRK2 and suggests that the overall mechanisms of Roc-COR signalling are conserved in all Roco proteins, although the details of the signal transduction as well as the final physiological outcomes could differ. Further research is required to establish whether or not the PD mutations in LRRK2 affect its nucleotide-induced conformational cycle.

Our HDX-MS experiments also revealed that distinct regions within the full-length CtRoco protein exist in multiple conformations at various times within the GTPase cycle (Supplementary Fig. S7). Two of these regions involve the P loop and switch I (residues 447-488) and the G4 loop (residues 566-588), while the last involves the 911-933 peptide within the C-COR domain, which is a critical dimerization surface. As mentioned previously, the P loop has the ability to occlude the nucleotide binding site (shown in Fig 1C). Given the importance of these regions in regulating the CtRoco protein, it is tempting to hypothesize that these multiple conformations play a functional role and that the conformational isomerism displayed in these regions could be potentially critical for controlling nucleotide binding and/or release as well as the concomitant monomerization/dimerization cycle. While conformational isomerism is often found in N-mers containing >4 subunits, it has been described in dimeric and tetrameric proteins (44,45). This conformational isomerism creates an ensemble of slowly interconverting structures, some of which are more or less resistant to nucleotide binding and/or to monomerization. Such behavior was called “deterministic” (45), to differentiate it from the stochastic behavior exhibited by most proteins. While there is a precedent for deterministic behavior, it is still a relatively rare phenomenon. Proteins exhibiting deterministic behavior exist at slightly different points along a given pathway or attain an ensemble of structures, rather than existing in one single point in equilibrium, or single structure. In our case, one potential scenario could be that the observed conformational isomerism might assist in releasing GDP after GTP hydrolysis. Most GTPases require guanine nucleotide exchange factors (GEFs) to destabilize the nucleotide binding site and release GDP (46,47). Roco family proteins might mimic this process in the absence of external GEFs by creating a nucleotide binding site able to achieve multiple conformations,

assisting the GDP release process. The consequence and potential drawback of having a nucleotide binding site capable of conformational isomerism would be lowered nucleotide affinities, as is observed for the Roco proteins (48). However, the benefit is that conformational isomerism could allow Roco proteins to cycle without the requirement for GEFs. We therefore could envision a hypothetical scenario in which dimerization of the nucleotide-free protein may slow the binding of GTP. The ensembles which have lost GDP and bind GTP will become activated and shift toward a monomeric structure. After the γ-phosphate is hydrolyzed, the protein shifts toward dimerization and the conformational isomerism destabilizes the nucleotide binding site enough that GDP is lost. This cycles the protein back to its nucleotide-free form and shifts the equilibrium further toward dimerization, where the combination of high cellular GTP concentrations and flexibility in the nucleotide binding site allow GTP to bind and the cycle to repeat.

Data Deposition

The atomic co-ordinates and structure factors have been deposited in the Protein Data Bank (PDB) with PDB ID code 6HLU.

Abbreviations

COR, C-terminal of Roc; CtRoco, Roco protein from *Chlorobium tepidum*; EM, electron microscopy; GDP, guanosine-5'-diphosphate; GppNHp, Guanosine-5'-[(β,γ)-imido]triphosphate; GTP, guanosine-5'-triphosphate; HDX-MS, hydrogen-deuterium exchange coupled to mass spectrometry; LRR, leucine-rich repeat; LRRK2, leucine-rich repeat kinase 2; MALS, multi-angle light scattering; MM, molecular mass; NF, nucleotide-free; PD, Parkinson's disease; Roc, Ras of complex protein; SAXS, small-angle X-ray scattering, SEC, size exclusion chromatography

Author contributions

ED, JL and WV designed the experiments. ED solved the crystal structure. ED and WV performed the SEC-SAXS analysis. ED and ML purified the proteins. RKS, ED and RG performed the SEC-MALS analysis. JL performed the HDX-MS experiments. ED, JS, AK, JL and WV interpreted the data. JL and WV wrote the manuscript. All the authors read and edited the manuscript.

Funding

This work was supported by the Fonds voor Wetenschappelijk Onderzoek (M.L., W.V.), a Strategic Research Program Financing of the VUB (W.V.), The Michael J. Fox Foundation for Parkinson's Research (R.K.S., A.K., W.V.), the Hercules foundation (W.V.), BioStruct-X by the European Community's Seventh Framework Programme (W.V.).

Acknowledgements

We would like to thank the staff at the beamline Proxima 2 of Soleil (France) and BM29 at ESRF (France) for assistance during data collection and the mass spectrometry facility at MPI-CBG for their support.

Declaration of Interests

The authors declare no competing interests.

References

1. Bosgraaf L, Van Haastert PJM. (2003) Roc, a Ras/GTPase domain in complex proteins. *Biochimica et Biophysica Acta - Molecular Cell Research.* **1643**, 5–10.

2. Marín I. (2006) The Parkinson disease gene LRRK2: Evolutionary and structural insights. *Mol Biol Evol.* **23**, 2423–33.

3. Marín I. (2008) Ancient origin of the parkinson disease gene LRRK2. *J Mol Evol.* **67**, 41–50.

4. Marín I, van Egmond WN, van Haastert PJM. (2008) The Roco protein family: a functional perspective. *FASEB J.* **22**, 3103–10.

5. Gotthardt K, Weyand M, Kortholt A, Van Haastert PJM, Wittinghofer (2008) A. Structure of the Roc-COR domain tandem of C. tepidum, a prokaryotic homologue of the human LRRK2 Parkinson kinase. *EMBO J.* **27**, 2239–49.

6. Terheyden S, Ho FY, Gilsbach BK, Wittinghofer A, Kortholt A. (2015) Revisiting the Roco G protein cycle. *Biochem J.* **465**, 139–47.

7. Gilsbach BK, Kortholt A. (2014) Structural biology of the LRRK2 GTPase and kinase domains: implications for regulation. *Front Mol Neurosci* **7**, 32.

8. Dihanich S. (2012) MASL1: a neglected ROCO protein. *Biochem Soc Trans* **40**, 1090–4.

9. Dihanich S, Civiero L, Manzoni C, Mamais A, Bandopadhyay R, Greggio E, et al. (2014) GTP binding controls complex formation by the human ROCO protein MASL1. *FEBS J.* **281**, 261–74.

10. Civiero L, Dihanich S, Lewis P a., Greggio E. (2014) Genetic, structural, and molecular insights into the function of RAS of complex proteins domains. *Chem Biol* **21**, 809–18.

11. Bialik S, Kimchi A. (2006) The death-associated protein kinases: structure, function, and beyond. *Annu Rev Biochem.* **75**, 189–210.

12. Carlessi R, Levin-Salomon V, Ciprut S, Bialik S, Berissi H, Albeck S, et al. (2011) GTP binding to the ROC domain of DAP-kinase regulates its function through intramolecular signalling. *EMBO Rep*. **12**, 917–23.

13. Jebelli JD, Dihanich S, Civiero L, Manzoni C, Greggio E, Lewis PA. (2012) GTP binding and intramolecular regulation by the ROC domain of Death Associated Protein Kinase 1. *Sci Rep.* **2**, 695.

14. Zimprich A, Biskup S, Leitner P, Lichtner P, Farrer M, Lincoln S, et al. (2004) Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. *Neuron.* **44**, 601–7.

15. Paisán-Ruíz C, Jain S, Evans EW, Gilks WP, Simón J, Van Der Brug M, et al. (2004) Cloning of the gene containing mutations that cause PARK8-linked Parkinson's disease. *Neuron* **44**, 595–600.

16. Lewis PA, Greggio E, Beilina A, Jain S, Baker A, Cookson MR. (2007) The R1441C mutation of LRRK2 disrupts GTP hydrolysis. *Biochem Biophys Res Commun.* **357**, 668–71.

17. West AB, Moore DJ, Choi C, Andrabi SA, Li X, Dikeman D, et al. (2007) Parkinson's disease-associated mutations in LRRK2 link enhanced GTP-binding and kinase activities to neuronal toxicity. *Hum Mol Genet.* **16**, 223–32.

18. Cookson MR. (2010) The role of leucine-rich repeat kinase 2 (LRRK2) in Parkinson's disease. *Nat Rev Neurosci.* **11**, 791–7.

19. Liao J, Wu C-X, Burlak C, Zhang S, Sahm H, Wang M, et al. (2014) Parkinson disease-associated mutation R1441H in LRRK2 prolongs the "active state" of its GTPase domain. *Proc Natl Acad Sci U S A* **111**, 4055–60.

20. West AB, Moore DJ, Biskup S, Bugayenko A, Smith WW, Ross CA, et al. (2005) Parkinson ' s disease-associated mutations in leucine-rich repeat kinase 2 augment kinase activity. *Proc Natl Acad Sci U S A*. **102**, 16842–7.

21. Greggio E, Jain S, Kingsbury A, Bandopadhyay R, Lewis P, Kaganovich A, et al. (2006) Kinase activity is required for the toxic effects of mutant LRRK2/dardarin. *Neurobiol Dis.* **23**, 329–41.

22. Gloeckner CJ, Kinkl N, Schumacher A, Braun RJ, O'Neill E, Meitinger T, et al. (2006) The Parkinson disease causing LRRK2 mutation I2020T is associated with increased kinase activity. *Hum Mol Genet*. **15**, 223–32.

23. Deng J, Lewis PA, Greggio E, Sluch E, Beilina A, Cookson MR. (2008) Structure of the ROC domain from the Parkinson's disease-associated leucine-rich repeat kinase 2 reveals a dimeric GTPase. *Proc Natl Acad Sci* **105**, 1499–504.

24. Guaitoli G, Raimondi F, Gilsbach BK, Gómez-Llorente Y, Deyaert E, Renzi F, et al. (2016) Structural model of the dimeric Parkinson's protein LRRK2 reveals a compact architecture involving distant interdomain contacts. *Proc Natl Acad Sci* **113**, E4357–66.

25. Deyaert E, Wauters L, Guaitoli G, Konijnenberg A, Leemans M, Terheyden S, et al. (2017) A homologue of the Parkinson's disease-associated protein LRRK2 undergoes a monomer-dimer transition during GTP turnover. *Nat Commun*. **8**, 1008.

26. Deyaert E, Kortholt A, Versées W. (2017) The LRR-Roc-COR module of the Chlorobium tepidum Roco protein: Crystallization and X-ray crystallographic analysis. A*cta Crystallogr Sect Struct Biol Commun.* **73**, 520–4.

27. Kabsch W. (2010) Xds. *Acta Crystallogr Sect D Biol Crystallogr.* **66**, 125–32.

28. Evans PR, Murshudov GN. (2013) How good are my data and what is the resolution? *Acta Crystallogr Sect D Biol Crystallogr.* **69**, 1204–14.

29. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. (2007) Phaser crystallographic software. *J Appl Crystallogr.* **40,** 658–74.

30. Afonine P V., Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, et al. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr Sect D Biol Crystallogr.* **68**, 352–67.

31. Shkumatov A V., Strelkov S V. (2015) DATASW, a tool for HPLC-SAXS data analysis. *Acta Crystallogr Sect D Biol Crystallogr.* **71**, 1347–50.

32. Svergun D, Barberato C, Koch MH. (1995) CRYSOL - A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr.* **28**, 768–73.

33. He W, Bai G, Zhou H, Wei N, White NM, Lauer J, et al. (2015) CMT2D neuropathy is linked to the neomorphic binding activity of glycyl-tRNA synthetase. *Nature* **526**, 710–4.

34. Mayne L, Kan ZY, Sevugan Chetty P, Ricciuti A, Walters BT, Englander SW. (2011) Many overlapping peptides for protein hydrogen exchange experiments by the fragment separation-mass spectrometry method. *J Am Soc Mass Spectrom.* **22**, 1898–905.

35. Walters BT, Ricciuti A, Mayne L, Englander SW. (2012) Minimizing back exchange in the hydrogen exchange-mass spectrometry experiment. *J Am Soc Mass Spectrom.* **23**, 2132–9.

36. Pascal BD, Willis S, Lauer JL, Landgraf RR, West GM, Marciano D, et al. (2012) HDX workbench: software for the analysis of H/D exchange MS data. *J Am Soc Mass Spectrom* **23**, 1512–21.

37. Wittinghofer A, Vetter IR. (2011) Structure-Function Relationships of the G Domain, a Canonical Switch Motif. *Annu Rev Biochem* **80**,943–71.

38. Verstraeten N, Fauvart M, Versees W, Michiels J. (2011) The Universally Conserved Prokaryotic GTPases. *Microbiol Mol Biol Rev* **75**, 507–42.

39. Pai EF, Kabsch W, Krengel U, Holmes KC, John J, Wittinghofer A. (1989) Structure of the guanine-nucleotide-binding domain of the Ha-ras oncogene product p21 in the triphosphate conformation. *Nature* **341**, 209–14.

40. Krissinel E, Henrick K. (2007) Inference of Macromolecular Assemblies from Crystalline State. *J Mol Biol.* **372**, 774–97.

41. Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, et al. (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–50.

42. Monfrini E, Di Fonzo A. (2017) Leucine-rich repeat kinase (LRRK2) genetics and parkinson's disease. *Advances in Neurobiology* **14**, 3–30.

43. Nixon-Abell J, Berwick DC, Grannó S, Spain VA, Blackstone C, Harvey K. (2016) Protective LRRK2 R1398H Variant Enhances GTPase and Wnt Signaling Activity. *Front Mol Neurosci* **9**, 18

44. Botelho MG, Rietveld AWM, Ferreira ST. (2006) Long-lived conformational isomerism of protein dimers: The role of the free energy of subunit association. *Biophys J.* **91**, 2826–32.

45. Erijman L, Weber G. (1991) Oligomeric Protein Associations: Transition from Stochastic to Deterministic Equilibrium. *Biochemistry* **30**, 1595–9.

46. Cherfils J, Zeghouf M. (2013) Regulation of Small GTPases by GEFs, GAPs, and GDIs. *Physiol Rev* **93**, 269–309.

47. Bos J, Rehmann H, Wittinghofer A. (2007) GEFs and GAPs : Critical Elements in the Control of Small G Proteins. *Cell* **129**, 865–77.

48. Wauters L, Terheyden S, Gilsbach BK, Leemans M, Athanasopoulos PS, Guaitoli G, et al. (2018) Biochemical and kinetic properties of the complex Roco G-protein cycle. *Biol Chem.* **399**, 1447-1456.

Tables

Table 1: Data collection and refinement statistics [a,b]	
Data Collection	
Diffraction source	PROXIMA 2, SOLEIL
Wavelength (Å)	0.9801
Detector	EIGER X 9M
Rotation range per image (°)	0.1
Total rotation range (°)	120
Space group	$P2_12_12_1$
a, b, c (Å)	95.6, 129.8, 179.5
α, β, γ (°)	90°, 90°, 90°
Resolution range (Å)	47.8 - 3.3 (3.4 - 3.3)
Total number of reflections	155879 (13110)
No. of unique reflections	34302 (3210)
Completeness	0.99 (0.96)
Multiplicity	4.5 (4.1)
Mean I/sigma(I)	9.3 (1.5)
R_{meas}	0.158 (1.103)
CC 1/2	0.995 (0.539)
Refinement	
R_{work} / R_{free}	0.2305 / 0.2813
No. atoms	13680
Average B-factor, all atoms (Å²)	87
Wilson B-factor (Å²)	85
Bond length RMS (deviation) (Å)	0.011
Bond angle RMS (deviation) (°)	0.901
Ramachandran favoured / allowed / outliers (%)	89 / 10 / 1
PDB code	6HLU

[a] **Data collection parameters are also published in** (26)
[b] **Values in parentheses are for the highest resolution shell**

Figure Legends

Figure 1: Structure of the LRR-Roc-COR domain construct of CtRoco. (A) Crystal structure of the CtLRR-Roc-COR homodimer in its nucleotide-free state, shown in two orientations. The LRR domain is colored yellow, the Roc domain red, the N-COR cyan, the C-COR blue and the connecting regions orange. The corresponding domains of the B-subunit are colored in a paler shade. **(B)** Comparison using CRYSOL of the theoretical scattering curve derived from the dimeric CtLRR-Roc-COR crystal structure (red line) with the experimental solution scattering curve of CtLRR-Roc-COR in the absence of nucleotides (black dots). **(C)** Zoom in on the Roc domains of the CtLRR-Roc-COR dimer crystal structure. The Roc domains from both subunits are colored dark and light grey respectively, with the most important regions colored as follows: P-loop: yellow, switch I: cyan, switch II: green, Roc dimerization loop: magenta, G4 loop: orange, G5 loop: salmon. The GppNHp molecules, that were placed in the active site by superposition on the structure of Ras in complex with GppNHp (PDB 5P21), are shown in transparent stick representation with carbon atoms colored green. Placement of GppNHp in the nucleotide-binding site would lead to steric clashes with the P loop due to the restrained conformation of the P loop in the dimer interface.

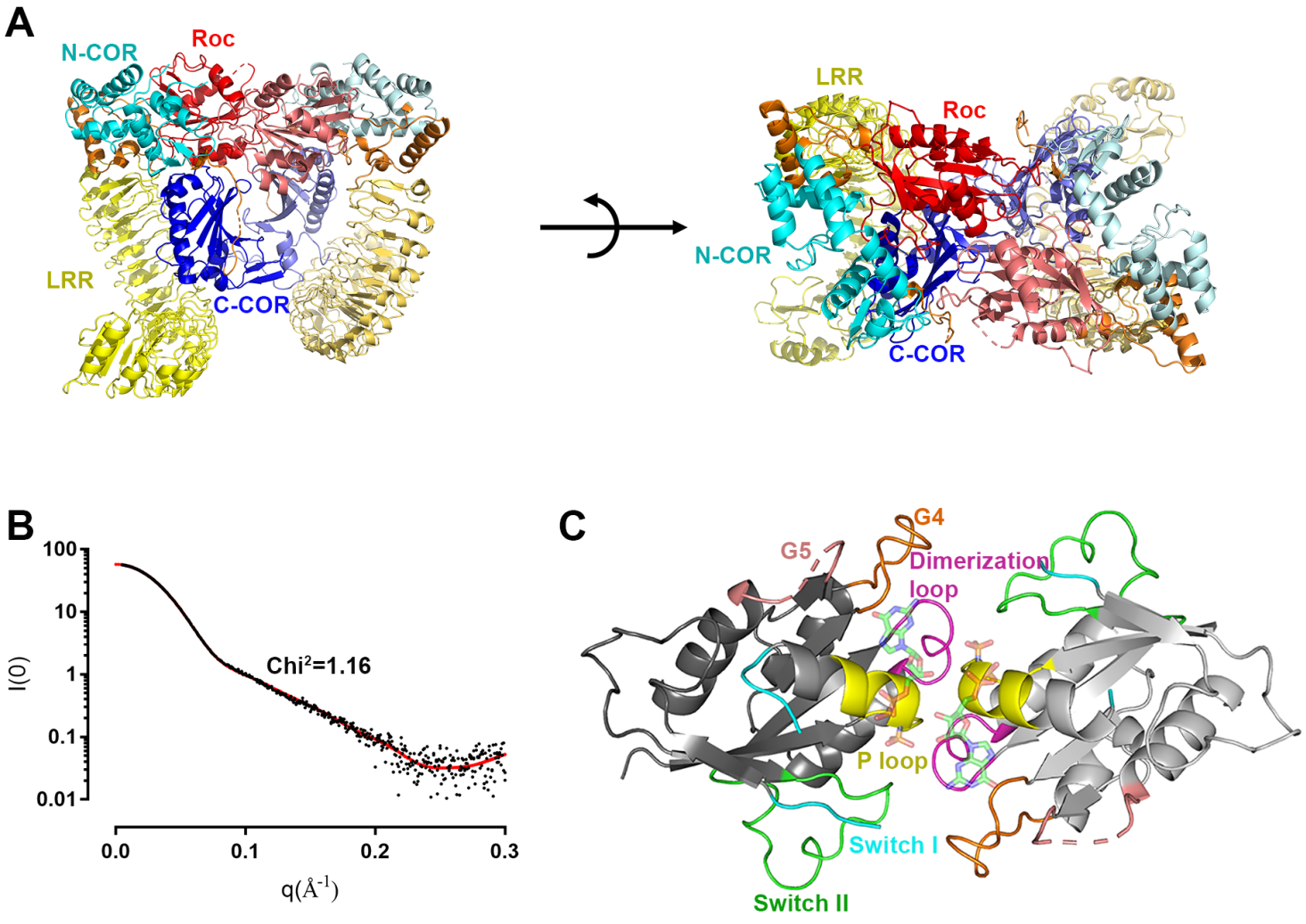
Figure 2: Dimer interface of CtLRR-Roc-COR. (A) CtLRR-Roc-COR dimer interface mapped on an accessible surface representation of the A subunit of the crystal structure, with the B subunit shown in coil representation (the LRR of the B subunit is not shown for clarity). The major interaction surfaces are colored according to the participating structural elements, with residues from the N-COR colored cyan, the C-COR colored blue and, within the Roc domain, residues from the P-loop yellow, from the switch II green, from the Roc dimerization loop magenta and from the G4 loop orange. **(B)** Schematic representation of the CtLRR-Roc-COR dimer interactions. Polar interactions (salt bridges and hydrogen bonds) are indicated by green lines, non-polar interactions by brown lines. **(C)** Zoom in on the Roc-Roc interface emphasizing the interactions between the Roc dimerization loop (magenta) of the A subunit and the P-loop (yellow), switch II (green) and Roc dimerization loop (magenta) of the B subunit. Polar interactions together with the interatomic distances are shown by sand-colored dotted lines.

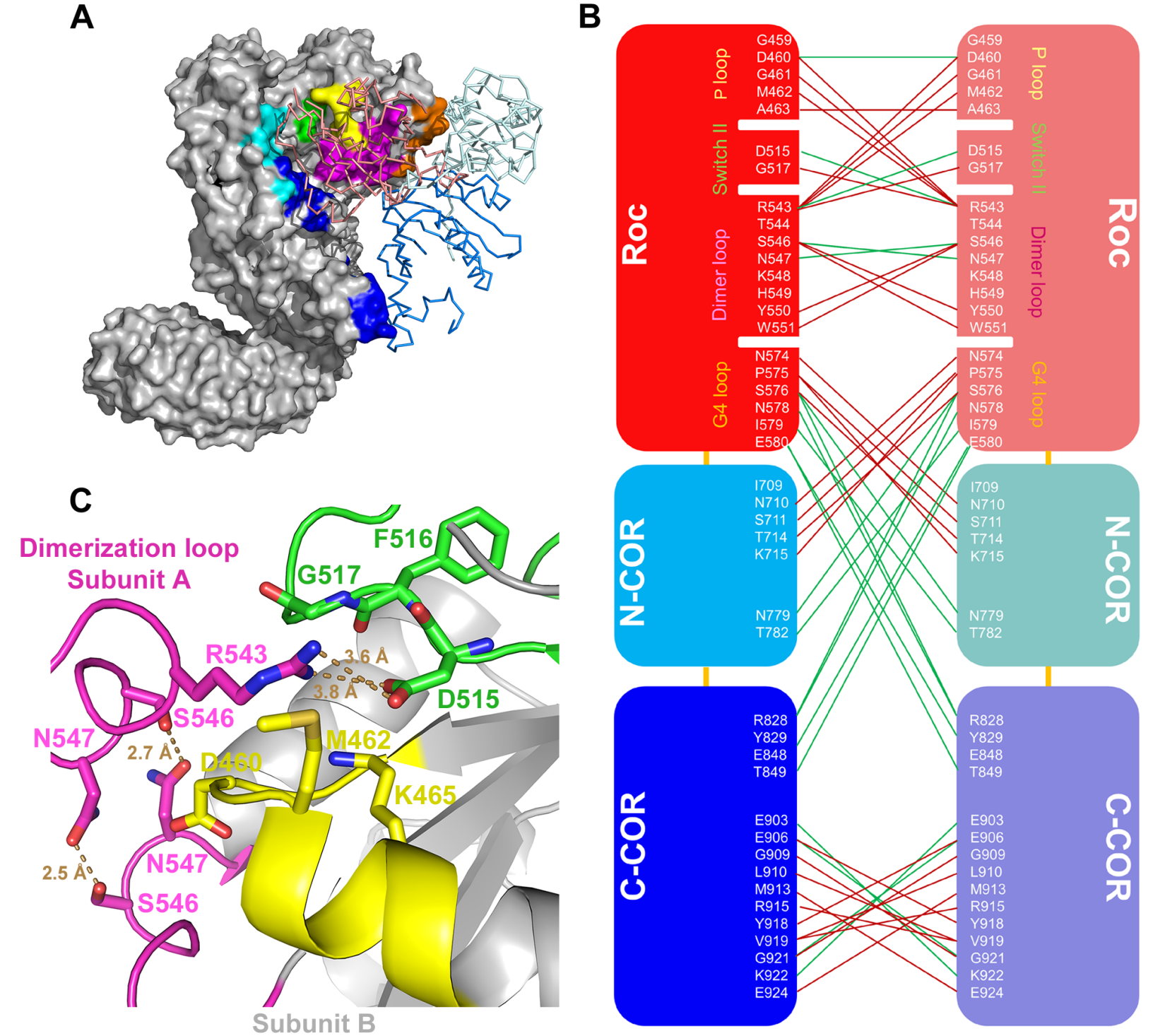
Figure 3: Switch II takes a central position at the domain interfaces. (A) Zoom in on the interaction interface between the C-terminal part of switch II (green), the LRR domain (yellow), the hinge region connecting the LRR to the Roc domain (orange) and the linker region connecting the Roc domain with the N-COR domain (salmon). Polar interactions together with the interatomic distances are shown by sand-colored dotted lines. The LRR-Roc hinge region spans the conserved ^{414}PLxxPPPE^{421} motif. Residue Arg532 links the switch II loop to the regions connecting the LRR to the Roc and the Roc to the N-COR domains. **(B)** Zoom in on the interactions of the N-terminal part of switch II (green) and the α3 helix of the Roc domain (grey) with the N-COR (cyan) and C-COR (blue) domains. The boxed residues His554, Tyr558 and Tyr804 align with Asn1437, Arg1441 and Tyr1699 in human LRRK2, which have been found mutated in PD.

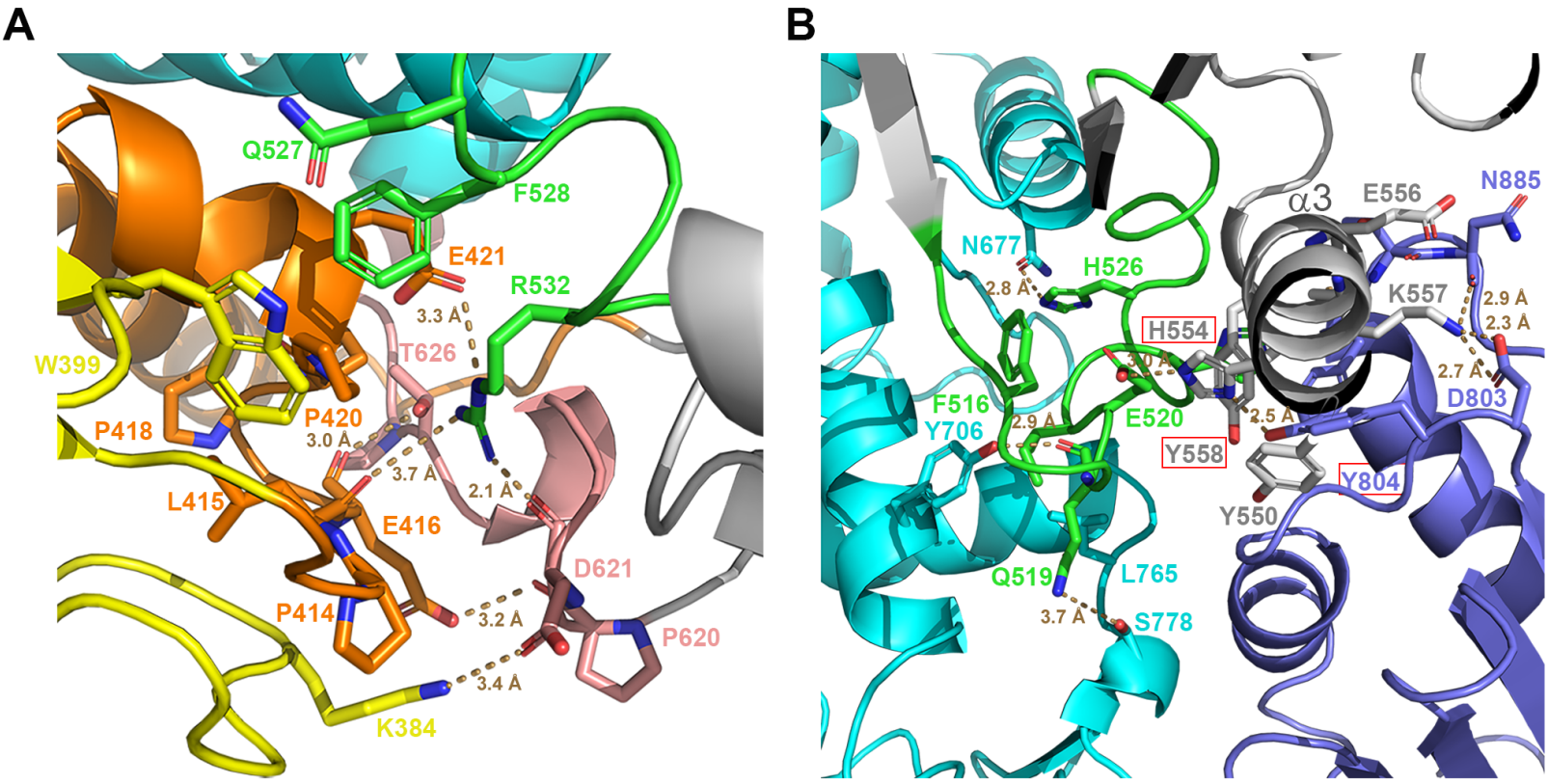
Figure 4: Nucleotide-induced conformational changes assessed by HDX-MS and mapped on the CtLRR-Roc-COR crystal structure. (A) HDX-MS data overlaid onto the structure of nucleotide-free CtLRR-Roc-COR. The left panel shows the alterations upon binding GppNHp. The middle panel

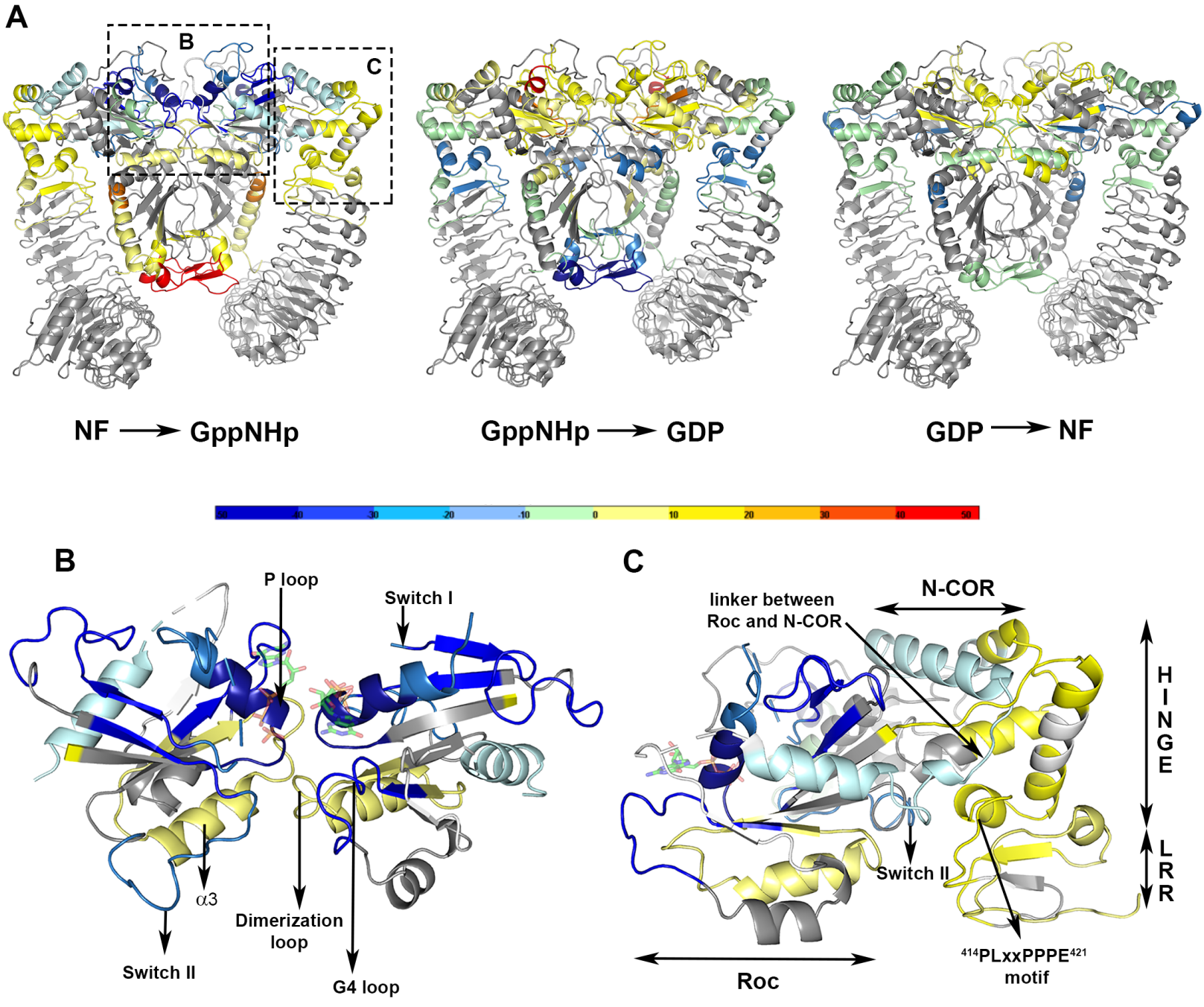
approximates the changes occurring during GTP hydrolysis by comparing GppNHp-bound protein to GDP-bound protein. The right panel shows the changes occurring upon GDP release. In each case, negative values (blue and green) are indicative of reduced deuterium uptake in the region. Positive values (red, orange, and yellow) correspond to enhanced deuterium uptake in a given region. **(B)** Zoomed in view on the left panel (GppNHp binding) of (A) highlighting important regions within the Roc-Roc dimer interface. **(C)** Zoomed in view on the left panel (GppNHp binding) of (A) highlighting important regions around the LRR – Roc - N-COR interface. In panels (B) and (C) a GppNHp molecule is modelled in and shown as transparent green sticks to indicate the position of the nucleotide binding site. Note that this position of GppNHp results in steric clashes with P-loop residues.

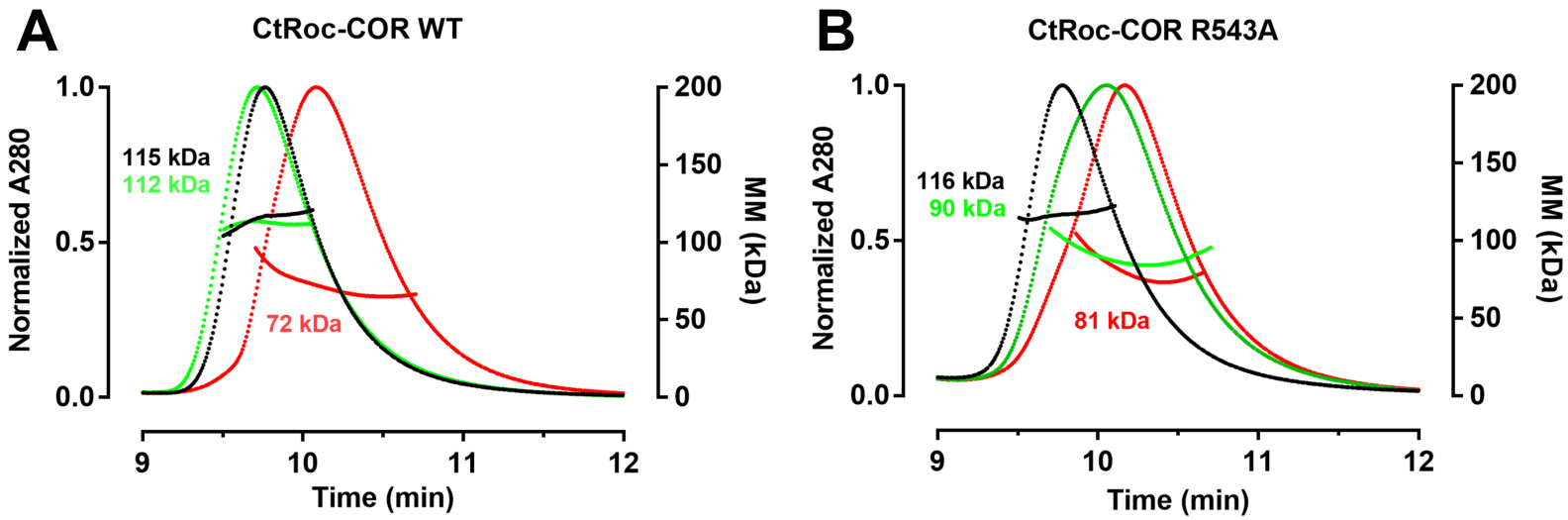
Figure 5: SEC-MALS analysis of the role of the Roc dimerization loop residue Arg543 in nucleotide-dependent CtRoc-COR homodimerization. (A) SEC-MALS analysis of the CtRoc-COR WT protein in absence (black) and presence of GDP (green) and GppNHp (red). **(B)** SEC-MALS analysis of the CtRoc-COR R543A mutant protein in absence (black) and presence of GDP (green) and GppNHp (red).


A
N-COR
Roc
LRR
C-COR
LRR
Roc
N-COR
C-COR
B
100
10
1
0.1
0.01
I(0)
Chi²=1.16
0.0
0.1
0.2
0.3
q(Å⁻¹)
C
G4
G5
Dimerization
loop
P loop
Switch I
Switch II


A
B
C
Dimerization loop
Subunit A
F516
G517
R543
3.6 Å
3.8 Å
D515
S546
N547
2.7 Å
D460
M462
K465
2.5 Å
N547
S546
Subunit B
Roc
P loop
Switch II
Dimer loop
G4 loop
G459
D460
G461
M462
A463
D515
G517
R543
T544
S546
N547
K548
H549
Y550
W551
N574
P575
S576
N578
I579
E580
N-COR
I709
N710
S711
T714
K715
N779
T782
C-COR
R828
Y829
E848
T849
E903
E906
G909
L910
M913
R915
Y918
V919
G921
K922
E924


A
Q527
F528
E421
R532
3.3 Å
T626
W399
P420
3.0 Å
P418
3.7 Å
2.1 Å
L415
E416
D621
P414
3.2 Å
P620
3.4 Å
K384
B
α3
E556
N885
N677
H526
2.8 Å
K557
2.9 Å
2.3 Å
H554
3.0 Å
2.7 Å
D803
F516
Y706
2.9 Å
E520
2.5 Å
Y558
Y804
Y550
Q519
L765
3.7 Å
S778


A
B
C
NF → GppNHp
GppNHp → GDP
GDP → NF
-50
-40
-30
-20
-10
0
10
20
30
40
50
B
P loop
Switch I
α3
Dimerization loop
Switch II
G4 loop
C
linker between Roc and N-COR
N-COR
HINGE
LRR
Switch II
Roc
414PLxxPPPE421 motif


A
CtRoc-COR WT
Normalized A280
MM (kDa)
115 kDa
112 kDa
72 kDa
Time (min)
B
CtRoc-COR R543A
116 kDa
90 kDa
81 kDa